THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans

Citation for published version:
Wells, GR, Weichmann, F, Colvin, D, Sloan, KE, Kudla, G, Tollervey, D, Watkins, NJ & Schneider, C 2016, 'The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans', *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw213

Digital Object Identifier (DOI):
10.1093/nar/gkw213

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Nucleic Acids Research

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 04. May. 2023

The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans

Graeme R. Wells[1], *Franziska Weichmann*[1], *David Colvin*[1], *Katherine E. Sloan*[1], *Grzegorz Kudla*[2,3], *David Tollervey*[2], *Nicholas J. Watkins*[1] and *Claudia Schneider*[1,] *

[1]Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK
[2]Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh, EH9 3JR, UK
[3]MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

* To whom correspondence should be addressed.

Claudia Schneider, PhD
Royal Society University Research Fellow
Institute for Cell and Molecular Biosciences
Newcastle University
Framlington Place
Newcastle upon Tyne
NE2 4HH, UK
Telephone: +44 (0) 191 208 7708
Fax: +44 (0) 191 208 7424
Email: Claudia.Schneider@ncl.ac.uk

Present Address:

Franziska Weichmann: Lehrstuhl für Biochemie I, Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, Regensburg, 93053, Germany

Katherine E. Sloan: Universitätsmedizin Göttingen, Institut für Molekularbiologie, Göttingen, 37073, Germany

Running Title: Endonucleases in yeast and human 18S rRNA release

Manuscript information:

Main text: ~9 printed pages (7440 words and 5 figures)
Supplementary Material: 4 tables and 6 figures

ABSTRACT

During ribosomal RNA (rRNA) maturation, cleavages at defined sites separate the mature rRNAs from spacer regions, but the identities of several enzymes required for 18S rRNA release remain unknown. PilT N-terminus (PIN) domain proteins are frequently endonucleases and the PIN domain protein Utp24 is essential for early cleavages at three pre-rRNA sites in yeast (A0, A1 and A2) and humans (A0, 1 and 2a). In yeast, A1 is cleaved prior to A2 and both cleavages require base-pairing by the U3 snoRNA to the central pseudoknot elements of the 18S rRNA. We found that yeast Utp24 UV-crosslinked *in vivo* to U3 and the pseudoknot, placing Utp24 close to cleavage at site A1. Yeast and human Utp24 proteins exhibited *in vitro* endonuclease activity on an RNA substrate containing yeast site A2. Moreover, an intact PIN domain in human UTP24 was required for accurate cleavages at sites 1 and 2a *in vivo,* whereas mutation of another potential site 2a endonuclease, RCL1, did not affect 18S production. We propose that Utp24 cleaves sites A1/1 and A2/2a in yeast and human cells.

Keywords: endonuclease cleavage/PIN domain/ribosome biogenesis/RNA cyclase

INTRODUCTION

The eukaryotic rRNAs are processed from the 35S (*Saccharomyces cerevisiae*) or 47S (*Homo sapiens*) rRNA precursors (pre-rRNAs) by initial endonucleolytic cleavages and subsequent exonucleolytic trimming, with concomitant removal of external (5'-ETS, 3'-ETS) and internal (ITS1, ITS2) transcribed spacer sequences (Figures 1 and S1) (reviewed in (1)).

Early pre-rRNA cleavages, at sites called A0, A1 and A2 in yeast and A0, A1/1 and 2a/E in humans, are important for 18S rRNA maturation. These events require a large ribonucleoprotein complex, the small subunit (SSU) processome (or 90S pre-ribosome) (reviewed in (2)). A key component of the SSU processome is the U3 small nucleolar (sno)RNA, which base-pairs with the 5'-ETS and 18S rRNA elements to chaperone the formation of the conserved pseudoknot, a key structural feature of the 40S ribosomal subunit (reviewed in (2,3)).

Loss of many different SSU processome components blocks pre-rRNA cleavage at sites A0, A1/1 and A2/2a, making it unclear which factor is the active nuclease. However, two evolutionarily conserved factors feature protein domains linked to RNA processing. Yeast Utp24/Fcf1 and human UTP24 harbor PIN (PilT N-terminus) endonuclease domains (4), whereas yeast Rcl1 and human RCL1 exhibit RNA cyclase-like protein folds (5). A PIN domain is also found in the endonuclease Nob1, which catalyzes the final cytoplasmic maturation step at the 3'-end of the 18S rRNA (site D) in yeast (6-8).

Yeast Utp24 and Rcl1 are essential for growth and conditional depletion of either protein inhibits A0, A1 and A2 cleavage (4,5,9). However, mutation of the PIN domain of Utp24 specifically inhibited cleavage at sites A1 and A2, while cleavage at site A0 was unaffected (4), and the mutant was dominant negative when expressed together with intact Utp24. These observations suggested that the presence of Utp24 is required for SSU processome function in A0-A2 cleavage, whereas the PIN domain harbors the catalytic activity for A1 and A2 cleavage (4), but nuclease activity was not demonstrated. Mutation of Rcl1 also inhibited site A2 cleavage, with less effect at A0 and A1 (5,9,10), and recombinant Rcl1 was reported to cleave pre-rRNA transcripts containing yeast site A2, consistent with direct endonuclease activity (9).

In mammals, the equivalent pre-rRNA cleavages also require the presence of UTP24 (FCF1 in mouse) and RCL1 (11-14) and the putative catalytic activity of human UTP24 was very recently linked to site 1 cleavage (15). However, the catalytic roles of UTP24 and RCL1 with respect to 2a cleavage have not been established.

In yeast, mutational and kinetic analyses indicate that processing at sites A1 and A2 is tightly coupled and mainly co-transcriptional (16,17). In vertebrates, pre-rRNA processing appears to be largely posttranscriptional and A′ processing near the 5'-end of the 5'-ETS in humans, a cleavage site not present in yeast, appears to be the only co-transcriptional event (18). Cleavage at site 2a within the human ITS1 (equivalent to yeast A2) is part of a minor pathway in humans (Figure S1) (11,12,19). However, no precursors processed at site 2a but still containing 5'-ETS sequences are detected, indicating that A0, 1 and 2a cleavages are coupled within the SSU processome, as in yeast.

Here, we present a combination of *in vivo* and *in vitro* approaches to clarify the roles of Utp24 and Rcl1 in yeast and human ribosome biogenesis. *In vivo* RNA-protein crosslinking studies (CRAC) generated a transcriptome-wide RNA binding profile for yeast Utp24, which provides fresh insights into its function. We also performed *in vitro* assays using recombinant wild type and mutant Utp24 and Rcl1 proteins to assess their cleavage activity on pre-ribosomal RNA. Finally, we established RNAi-rescue systems in HEK293 cells to study the effect of presumably catalytically inactive UTP24 and RCL1 mutants on pre-rRNA cleavage in the human system.

MATERIALS AND METHODS

Yeast strains and methods

S. cerevisiae strains (Table S1) were constructed by standard methods (20). Cultures were grown at 30°C in medium containing 0.67% nitrogen base (Difco) and 2% glucose or 2% galactose. Strains for crosslinking studies expressed genomically encoded, C-terminal HTP-tagged proteins under the control of their endogenous promoter.

CRAC and data analysis

The CRAC method was performed as previously described (21) (22), see Figure S2A. To generate RNA-protein crosslinks, actively growing yeast cultures in SD medium (OD_{600} ~0.5) were UV-irradiated in a 1.2 m metal tube for 100 sec at 254 nm (22). Illumina sequencing data was aligned to the yeast genome using Novoalign (http://www.novocraft.com). Bioinformatics analyses were performed as described (23). The Illumina sequencing data from this publication have been submitted to the GEO database (http://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE75991.

Cloning and mutagenesis

The open reading frames for Rcl1, Bms1, Utp24 or UTP24 were amplified from yeast genomic DNA or human cDNA adding restriction sites (Table S2) and cloned into pET100 vectors (Invitrogen). The constructs were used for *in vitro* translation in the presence of [^{35}S] methionine (TNT, Promega) or subcloned into pGEX-6P1 vectors to express and purify GST-tagged recombinant proteins from *E.coli* using standard techniques. The C-terminally HTP-tagged Rcl1 gene was amplified by PCR from yeast genomic DNA (Strain Rcl1-HTP) adding restriction sites to the 5'-end of the gene and the 3'-end of the HTP-tag, respectively, and cloned into pRS316. The coding sequences of UTP24 and RCL1 were altered to make them resistant to the siRNAs (Table S3) used to deplete the endogenous mRNAs. These constructs (IDT) were amplified by PCR and cloned into the pcDNA5/FRT/TO vector (Invitrogen) containing 2x N-terminal FLAG tags under the control of a tetracycline-inducible promoter. Point mutations were generated by QuikChange site-directed mutagenesis using overlapping primers (Table S2) and verified by sequencing.

Cell Culture and RNAi

Constructs were transfected into Flp-In T-Rex HEK293 cells. Stably transfected cells were selected as described by the manufacturor (Invitrogen) and cultured according to standard protocols. Expression of exogenous proteins was induced by addition of tetracycline (UTP24; 1 mg/ml, RCL1; 0.01-0.1 mg/ml). Cells were transfected with siRNA duplexes (Table S3) using Lipofectamine RNAiMAX transfection reagent (Invitrogen) and harvested after 72h depletion.

RNA analysis

RNA was extracted from HEK293 cell pellets using TRI reagent (Sigma-Aldrich). For northern blot analysis, 2 µg of total RNA was separated on 1.2% glyoxal-agarose gels, transferred to nylon membrane and hybridized with 5'-end labeled oligonucleotide probes (Table S2). For primer extension analysis, 1 µg of total RNA was converted into cDNA using Superscript III (Invitrogen) and 5'-end labeled oligonucleotide probes and separated on 10% PAA/8M urea sequencing gels. Results were visualized using a PhosphorImager (Typhoon FLA9000; GE Healthcare). ImageQuant (GE Healthcare) was used to quantify northern blot data, which were normalized to levels of the 47S/45S pre-rRNAs.

In vitro RNA cleavage assay

Recombinant proteins were expressed and purified as described in (24), including 1 mM $MnCl_2$ in all media and buffers. Nuclease assays were performed in 10 mM Tris/HCl pH 7.6, 100 mM NaCl, 2 mM dithiothreitol, 100 µg ml^{-1} bovine serum albumin, 0.8 unit $µl^{-1}$ RNasin, 4.5% glycerol, 0.05% Tween20, 10 ng $µl^{-1}$ *E.coli* tRNA and 5 mM $MnCl_2$. 10 µl reactions containing ~20 pmol of protein were pre-incubated for 5 min at 30°C. The only exception was for the experiment shown in Figure S4, for which detailed conditions are described in the supplementary figure legend S4. *In vitro* transcribed pre-rRNA substrates (0.125 pmol) containing yeast sites D and A2 (35S+2301-2844) or yeast sites A0 and A1 (35S+335-1146) was added and incubated for 1h at 30°C. Following proteinase K digestion (30 min at 37°C), the RNA was extracted, precipitated and analyzed by primer extension using probes yD-RT, yA2-RT or yA1-RT, respectively. Products were resolved on 10% PAA/8M urea sequencing gels and visualized by autoradiography.

Protein-protein interaction studies and purification of Rcl1-containing complexes

GST-bait proteins (~50 pmol) were immobilized on glutathione sepharose and incubated with [^{35}S] *in vitro* translates for 1h at 4°C in Buffer NB (20 mM Tris/HCl pH 7.6, 150 mM NaCl, 8.7% glycerol and 0.1% Tween20). The beads were washed five times with buffer NB. Retained proteins were separated by SDS-PAGE and visualized by autoradiography. Affinity-purification of Rcl1-HTP complexes (from 500 ml of yeast culture each) was performed essentially as described in (25). Complexes were purified on IgG sepharose at 150 mM NaCl and eluted by TEV protease cleavage. Co-purifying proteins and RNAs were analyzed by immunoblotting or northern blotting, respectively (Tables S4 and S2).

Yeast Immunofluorescence

Yeast strains expressing HA-tagged Rcl1 under control of the repressible *GAL10* promoter were transformed with plasmids encoding His_6-TEV-protein A (HTP) tagged forms of Rcl1 and $Rcl1_{RDK}$ and grown in minimal medium containing glucose for 6h to deplete the endogenous protein. Cells at mid-log phase were harvested and fixed in 4% formaldehyde in PBS (15 min at RT), washed with PBS and then incubated for 45 min at 30°C in buffer B (0.1 M potassium phosphate pH 7.5, 1 M sorbitol, 10 mM DTT) containing 50 U/ml Zymolase. Cells were washed and incubated in batch with the antibodies diluted in PBS + 5% milk (Table S4), washed with PBS containing DAPI (4',6'-diamidino-2-phenylindole) and mounted onto poly-lysine coated coverslips using Vectorshield. All images were obtained using a Zeiss Axiovert 200 microscope with Plan-Apochromat x100 1.4NA objective, Axiovision software and an Axiocam monochrome camera, and processed in Photoshop (Adobe).

For Peer Review

RESULTS

Yeast Utp24 crosslinks in close vicinity to site A1, to the U3 snoRNA and pre-rRNA elements required for pseudoknot formation

To dissect the roles of the two yeast candidate pre-rRNA endonucleases Rcl1 and Utp24, we applied *in vivo* RNA-protein crosslinking (CRAC) to identify their RNA binding sites (21) (Figures 1 and S2). C-terminal HTP-tagged (His$_6$-TEV-protA) Rcl1 and Utp24 were expressed from the chromosomal locus under control of the endogenous promoter. Rcl1-HTP supported wild-type growth, while the Utp24-HTP strain exhibited a mild growth defect. Actively growing cells were UV-irradiated as described (22) and RNA fragments crosslinked to the purified proteins were isolated and analyzed as outlined in Figure S2A. Protein recovery was verified by western analysis (Figure S2B). Rcl1 purified well, but crosslinked poorly, and no pre-rRNA target sequence was significantly enriched in every experiment. We were, however, able to reproducibly identify Utp24 RNA crosslinking sites in four independent CRAC experiments. Results from two representative datasets are presented in Figures 1 and S2.

Transcriptome-wide RNA binding profiles of Utp24 are shown in Figure 1A. The majority of reads (> 83%) identified in both Utp24 datasets were mapped to the (pre-)rRNA, with lower numbers of hits in snoRNAs. Hits were also recovered in mRNAs, but no individual mRNA emerged as a likely Utp24 target. Consistent with its known role in SSU biogenesis, Utp24 was predominately associated with sequences within the 18S rRNA (Figures 1B and C). Importantly, this crosslinking profile was not seen in a CRAC experiment performed with an untagged control strain (Figure 1B). The peak near the 3'-end of the 25S rRNA represents a common CRAC contaminant (21).

In CRAC, microdeletions and/or mutations are often introduced at the site of crosslinking during cDNA preparation and can be used to map precise protein-binding sites (26). Sites of microdeletion (Figures 1C and D) were seen at 18S+3, adjacent to cleavage site A1 (peak 3, see Figure S2C for higher resolution), around 18S+545 (peak 4) and around 18S+820 and 18S+895 within expansion segment 6 (ES6) (peak 5), close to the binding site for the snoRNA snR30, which is also required for cleavage at sites A0-A2 (27). The highest peak was located at 18S+1103 in the vicinity of the 3'-side of the central pseudoknot (peak 6). Sites of microdeletion were also identified in the flanking 5'-ETS (Figure 1E) around 35S+288 (peak 1) and 35S+460 (peak 2). Strikingly, these positions in the 5'-ETS correspond to the base-pairing sites for the U3 snoRNA 3'-hinge and 5'-hinge regions, respectively (reviewed in (3)).

Analysis of snoRNAs crosslinked to Utp24 (Figure 1F) revealed specific enrichment for the U3 snoRNA (encoded by the genes *SNR17A/B*). Utp24 predominately crosslinked to a U3 region (+15-90), which undergoes multiple interactions with the pre-rRNA. Significant peaks of microdeletions in U3 (Figures 1G and 1H) were seen at +45 and +52 (peaks b, 5'-hinge) and +68 (peak c, 3'-hinge) that base-pair to 5'-ETS regions at +480 and +280, respectively (reviewed in (2,3)). The highest peak of microdeletions (peak a) was at +24, a region of U3 predicted to base-pair with 18S+1140 on the 3'-side of the pseudoknot.

We conclude that Utp24 binds both to the U3 snoRNA and to the corresponding U3-binding sequences in the mature 18S rRNA and the 5'-ETS, interactions that are essential for pseudoknot formation and coupled cleavage of sites A0, A1 and A2. Moreover, the binding of Utp24 at 18S+3 is consistent with a direct role in cleavage at site A1, since the precise location of cleavage is partly defined with respect to this position (28).

Utp24 cleaves site A2 in a yeast pre-rRNA substrate *in vitro*

The predicted catalytic center within the Utp24 PIN domain is characterized by four acidic residues (DEDD), which are conserved in all kingdoms of life (29). We therefore assessed the *in vitro* activities of yeast Utp24 and human UTP24 in cleavage assays (Figures 2, S3 and S4). Wild type and PIN domain mutant proteins were expressed in *E. coli* with N-terminal tags containing GST. Double (D72N/D142N) catalytic site mutations in UTP24 were

based on yeast Utp24 mutations (D68N, D138N) that cause A1 and A2 cleavage defects *in vivo* (4). The N-terminal GST tag was removed by prescission protease (PP) cleavage following purification (Figure S3A).

The proteins were tested in nuclease assays on *in vitro* transcribed RNA containing yeast site A2 (Figure 2A) or A1 (Figure S3B). Assays were performed in the presence of 5 mM Mn^{2+}, which is required for the *in vitro* activity of other PIN domain nucleases (7,25,30). The RNA was analyzed by primer extension using a labeled oligonucleotide downstream of the cleavage site.

Wild type yeast and human Utp24 both cleaved the RNA substrate containing site A2, giving very similar products (Figure 2B, lanes 8 and 12). In contrast, no cleavage activity was observed for the predicted catalytically inactive Utp24 PIN mutants (Figure 2B, lanes 9, 10 and 13), while a different PIN domain endonuclease, Nob1, cleaved site D but not A2 (Figure S3D). Utp24-mediated cleavage did not occur exclusively at site A2, but also at several other places with the same sequence context (5'C/A3', where "/" equals the site of cleavage) around the *bona fide* cleavage site. In contrast, we did not detect clear cleavage at site A1 (Figure S3B, see Discussion). An evolutionarily conserved ACAC motif at site A2 was previously identified as a specificity feature for yeast A2 cleavage *in vivo* (31). Since the *in vitro* cleavage pattern was identical for yeast and human Utp24 in our assays, this appears to be a conserved feature of Utp24 PIN nuclease activity.

Yeast Rcl1 was also expressed and purified with a cleavable GST-PP tag (Figures 2 and S3A), but no *in vitro* cleavage activity was detected in the presence of 5 mM Mn^{2+} (Figure 2B, lane 11). In an effort to reproduce the reported nuclease activity of Rcl1 (9), we dialyzed the yeast Rcl1 protein, as well as wild type and mutant Utp24 proteins, into the buffer optimized for Rcl1 activity (K. Karbstein, personal communication) and performed the nuclease assay in the presence of 10 mM Mg^{2+} Under these conditions, apparent cleavage at site A2 was observed (Figure S4), as reported. However, cleavage was observed with each of the recombinant proteins, including the catalytically inactive Utp24D68N mutant, indicating that cleavage associated with addition of Utp24 was not specific. We were unable to express and purify a mutant of Rcl1 that is predicted to lack catalytic activity, and were therefore unable to determine whether this is also the case for Rcl1.

An intact PIN domain in human UTP24 is required for accurate cleavages at sites 1 and 2a, but not A0, *in vivo*

Yeast cleavage sites A0-A2 have direct counterparts in humans called sites A0, 1 and 2a (Figure S1). RNAi-mediated knockdown of human UTP24 caused 30S pre-rRNA accumulation, indicating that UTP24 is required for these cleavages (11,13). To investigate the putative catalytic role of UTP24, we established an RNAi-rescue system in HEK293 cells. Cells stably expressing the FLAG-tag (pcDNA5) or FLAG-tagged forms of either wild type UTP24, or UTP24 with single (D72N) or double (D72N/D142N) catalytic site mutations were generated (Figures 3 and S5). These carried silent mutations in the open reading frames rendering them resistant to RNAi-mediated depletion of the endogenous protein. The RNAi-resistant proteins were expressed at endogenous levels using a titratable *TET* promoter. Cells were transfected with either a siRNA specifically targeting endogenous UTP24 or a control siRNA targeting firefly luciferase (GL2) (32). After siRNA treatment for 72h, the expression of endogenous and FLAG-tagged UTP24 proteins was analyzed by immunoblotting (Figure S5A). The UTP24-specific siRNA significantly reduced endogenous protein levels, whereas the RNAi-resistant tagged UTP24 proteins were unaffected.

Total RNA was extracted from RNAi-treated cells and pre-rRNA processing analyzed by northern hybridization using probes complementary to the 5'-end of ITS1 ("h18SE") or downstream of the 2a cleavage site ("hITS1") (Figure 3B and S5B). Depletion of UTP24 resulted in 30S accumulation (Figure 3B, lane 2), and this phenotype was almost completely rescued by expression of the wild type protein (lane 4). Complementation with UTP24 D72N or D72N/D142N mutants caused accumulation of the 26S pre-rRNA (lanes 6 and 8), indicative of strongly reduced cleavage at sites 1 and 2a. No dominant negative phenotypes

were observed upon expression of the PIN mutants in the presence of the endogenous protein (lanes 5 and 7).

The 26S RNA accumulation indicated that the catalytic activity of human UTP24 is required for site 1 and 2a cleavage *in vivo*, while cleavage at site A0 is unaffected. However, 2a cleavage cannot readily be directly assessed because the 36S pre-rRNA and the excised ITS1 fragment, the only intermediates specific to the "minor" pathway (lighter shades in Figure 3A), are barely detectable in control cells. However, depletion of the 5'-3'-exonuclease XRN2 stimulates processing through this pathway (11,19,33). We therefore performed UTP24 knockdown and complementation in combination with XRN2 depletion. Pre-rRNA intermediates were detected by northern hybridization with probe hITS1 (Figure 3C), and quantified using a PhosphorImager. Levels were normalized to the 47S/45S pre-rRNA and plotted to compare the single XRN2 and UTP24/XRN2 double knockdowns (Figure 3D). Depletion of XRN2 alone caused strong accumulation of the 36S pre-rRNA and the ITS1 fragment accompanied by the appearance of the 30SL5' precursor (also referred to as 34S (12) (Figure 3C, lanes 1, 3, 5 and 7), reflecting an A' cleavage defect. Double-knockdown of UTP24 and XRN2 (lane 2) resulted in a strong decrease in the 36S precursor and the ITS1 fragment compared to the single XRN2 knockdown, while expression of the wild type UTP24 protein rescued the 2a processing defect (lane 4). Importantly, the 36S pre-rRNA and the excised ITS1 fragment were also severely reduced in double knockdown cells expressing UTP24 D72N (lane 6) or UTP24 D72N/D142N (lane 8), whereas 26S pre-rRNA levels were strongly increased.

Low levels of the 21S precursor were detected in UTP24 PIN mutant cells, suggesting that an alternative mechanism can generate the 18S rRNA 5'-end (site 1) in the absence of endonuclease cleavage. XRN2 degrades the 3'-fragment of the 5'-ETS ("ETS3") (13) and might also digest the 5'-ETS back to site 1 following A0 cleavage. To test this model, RNA was extracted from cells expressing UTP24 D72N, treated with a siRNA against endogenous UTP24, XRN2 or both, and analyzed by primer extension through site 1 (Figure 3E). The major stop at site 1 reflects mature 18S rRNA produced prior to the siRNA treatment. In cells treated with the UTP24 siRNA (lane 7), a weaker primer extension stop was visible two nucleotides downstream of site 1. This was significantly reduced in the UTP24/XRN2 double knockdown (lane 9) and absent when the catalytically inactive UTP24 was not expressed (Figure S5C). We speculate that in the absence of site 1 cleavage, XRN2 degrades the 5'-ETS region, but can be blocked by catalytically inactive UTP24 bound at the 5'-end of 18S, leading to the observed truncated product.

Our analyses indicate that point mutations in the PIN domain of UTP24 block cleavage at sites 1 and 2a in human pre-rRNA, while cleavages at sites A0 and A' are unaffected.

Mutation of the proposed pre-rRNA substrate binding site within human RCL1 does not affect 18S production or 2a cleavage

Yeast Rcl1 was also reported to cleave at site A2 (9), suggesting the possibility of redundant activities. To determine whether RCL1 participates in human site 2a cleavage, we established an RCL1 RNAi-rescue system in HEK293 cells, as described for UTP24 (Figures 4 and S6). Endogenous and FLAG-tagged RCL1 levels were monitored by immunoblotting (Figure S6A). RNAi-mediated depletion of RCL1 caused 30S accumulation (Figure 4A, lane 2), that was rescued by expression of RNAi-resistant FLAG-tagged wild type RCL1 (Figure 4A, lane 4).

In yeast Rcl1 an "RDK-AAA" mutant showed a strong A2 cleavage defect, which was proposed to be due to impaired pre-rRNA substrate binding (9). A stable cell line was constructed expressing RCL1 with the equivalent "RHK" residues (R329, H330, K332) changed to alanines. Following RCL1 depletion, $RCL1_{RHK\text{-}AAA}$ fully rescued the 30S phenotype, showing that integrity of the proposed RNA substrate-binding pocket is not required for human 18S maturation (Figure 4A, lane 6).

As noted above, XRN2 knockdown increases dependency on 2a cleavage, so the RCL1 knockdowns were repeated in combination with XRN2 depletion (Figure 4B). Following

RNAi-mediated knockdown of XRN2 and RCL1, wild type RCL1 and the $RCL1_{RHK}$ mutant again rescued the pre-rRNA processing phenotype to a similar extent (Figure 4B, lanes 4 and 6). Quantification of pre-rRNA levels normalized to the 47S/45S pre-rRNAs confirmed that the 2a-dependent pre-rRNA intermediates, the 36S pre-RNA and the ITS1 fragment, are restored to similar levels in cells expressing wild type and mutant RCL1, but severely reduced in cells expressing the FLAG-tag only (Figure 4C).

Yeast Rcl1 mutations impair interactions with Bms1, integration into the SSU processome and subcellular localization

The finding that the human equivalent of the yeast $Rcl1_{RDK}$ mutant supports pre-rRNA processing prompted us to re-examine the ribosome synthesis defect in yeast (Figure 5). Recent mutational analysis of Rcl1 based on the crystal structure of the yeast Rcl1-Bms1 dimer showed that mutation of R237, within the RDK motif, disrupts interaction with Bms1 (10). The Rcl1-Bms1 interaction is required for nuclear import of Rcl1 (34) and, consistent with this, an R237A mutation severely impaired the nucleolar localization of yeast Rcl1 (10).

Wild type Rcl1 and $Rcl1_{RDK}$ were generated by *in vitro* translation in the presence of [^{35}S] methionine. Incubation with recombinant GST-tagged $Bms1^{(aa1\text{-}705)}$ (Figure 5A) revealed a strong and reproducible interaction between the N-terminal region of Bms1 and the wild type Rcl1 protein, whereas no binding was detected for the $Rcl1_{RDK}$ mutant.

This finding suggested that $Rcl1_{RDK}$ might not be incorporated into the SSU processome *in vivo*. To test this, we generated a yeast strain expressing HA-tagged Rcl1 under control of the repressible *GAL10* promoter, allowing depletion of the endogenous Rcl1 protein. This strain was then transformed with plasmids encoding His_6-TEV-protein A (HTP) tagged forms of Rcl1 and $Rcl1_{RDK}$, or an empty plasmid (pRS316) and grown in minimal medium containing glucose. Rcl1-containing complexes were purified on IgG sepharose followed by TEV cleavage, and associated proteins and RNA were detected by western and northern blotting, respectively (Figure 5). This revealed that Rcl1 and $Rcl1_{RDK}$ were expressed and purified to similar levels. However, association of $Rcl1_{RDK}$ with the boxC/D snoRNP-specific protein Nop1 (Figure 5B) as well as the U3 snoRNA and several pre-rRNA processing intermediates (23S, 22S/21S, 20S) (Figure 5C) was severely reduced, confirming that the $Rcl1_{RDK}$ mutant fails to associate with the SSU processome. To support this finding, we performed immunofluorescence in cells expressing HTP-tagged Rcl1 and $Rcl1_{RDK}$ (Figure 5D). Wild type Rcl1 localized to the nucleolus as expected, whereas $Rcl1_{RDK}$ was mainly retained in the cytoplasm.

We conclude that pre-rRNA processing defects associated with yeast $Rcl1_{RDK}$ are likely due to impaired nuclear import and SSU processome formation, due to the loss of binding and recruitment via Bms1.

DISCUSSION

Here, we present data implicating Utp24 as the endonuclease responsible for early pre-rRNA cleavages at sites A1/1 and A2/2a that generate the major precursor to the 18S rRNA.

In vivo UV crosslinking was performed in actively growing cells to identify *bona fide* RNA targets for yeast Utp24 (Figures 1 and S2). Utp24 RNA binding was strongest in the region of the 18S rRNA central pseudoknot, a highly conserved structural unit that forms the core of the small ribosomal subunit. Additional crosslinking was observed close to site A1 and in the 5'-ETS pre-rRNA spacer region. Pseudoknot formation is guided by U3 and is required for A0, A1 and A2 cleavages (reviewed in (3)). Strikingly, Utp24 was recovered in association with the U3 snoRNA regions that base-pair with the pre-RNA as well as with the U3-binding sites in the 5'-ETS and the 18S elements forming the proximal and distal sides of the pseudoknot (see Figure 1H). These finding indicate a role for U3 in Utp24 recruitment to the A1 cleavage site and, potentially, a role for Utp24 in verification of U3-rRNA interactions.

Utp24 was also crosslinked to the eukaryotic expansion segment 6 (ES6) within the 18S rRNA central domain (Figure 1D). The ES6 region has recently emerged as a binding hub for many ribosome biogenesis factors including the RNA helicase Dhr1, which is required to dissociate U3 from the pre-rRNA following cleavages at sites A0 and A1, and preceding A2 cleavage (35). Utp24 and Dhr1 both crosslink to the same U3 snoRNA regions, which engage in pre-rRNA interactions (35). It is possible that sequential binding of both proteins to U3 might be important to guide structural rearrangements to promote U3 release and/or positioning of Utp24 for A1 and A2 cleavages.

We did not detect clear Utp24 crosslinking around site A2 in the ITS1 sequence. We speculate that Utp24 is recruited to stable docking points within the mature 18S rRNA sequence, and only transiently interacts with the cleavage sites during catalysis. The observed coupling of cleavage at sites A1 and A2 suggests that they are brought into proximity by transient changes in pre-ribosome structure, which remain to be determined.

In vitro endonuclease assays revealed that recombinant Utp24 proteins from yeast and human each show activity on a pre-rRNA substrate containing yeast site A2 (Figure 2). As for other PIN domain proteins (7,25,30), this activity required Mn^{2+} ions and the conserved metal-binding amino acids of the PIN domain. Interestingly, yeast and human Utp24 not only cleaved the authentic site A2 (5'-AAC/ACAC-3'), but also at other positions in a similar sequence context (Figure 2). The cleavage pattern is identical in both proteins and therefore likely specific to the conserved Utp24 PIN domain. Very low levels of cleavage at these positions around site A2 are also observed *in vivo* (31), but they are underrepresented compared to the *bona fide* A2 site. Bound transacting factors within the SSU processome therefore likely assist the specificity of Utp24-mediated cleavage *in vivo*. The recently mapped 2a site in the human pre-rRNA (5'-C/GAC/GC-3) (19) appears to be very different to yeast A2. However, cleavage takes place at 5'C/purine3' in both cases. We also tested RNA substrates including yeast site A1, but observed only low levels of cleavage in ACAC-rich regions within the 5'-ETS close to the U3 base-pairing sites (Figure S3B and data not shown). Notably, the sequence context of site A1 (5'-GU/UA-3') does not match yeast A2 or human 2a, highlighting the importance of additional factors and/or structural arrangements within the pre-rRNA that might be needed to facilitate Utp24-mediated cleavage at this site *in vivo*.

Consistent with yeast data, *in vivo* studies on human UTP24 showed that an intact UTP24 PIN domain was required for cleavage at the 5'-end of 18S (site 1) and at site 2a within the ITS1 spacer, but not at A0 or A' in the 5'-ETS (Figure 3). The requirement for UTP24 in site 1 cleavage and the presence of an alternative pathway to process to the 5'-end of 18S in the absence of endonuclease cleavage by UTP24 was recently reported (15). However, in contrast to our results, XRN2 was not implicated in 5'-3'-exonuclease trimming to generate the aberrant 5'-end of 18S. This discrepancy might reflect the use of stable shRNA expression by Tomecki et al. (15), which strongly impaired cellular viability and could result in secondary effects. Notably, inspection of published northern analyses reveals phenotypes that link UTP24 to 2a cleavage (Figure S10 in (15)).

Both Rcl1 and Utp24 have previously been proposed to cleave site A2 in yeast (4,5,9). A2 cleavage can occur either co- or post-transcriptionally, so two separate enzymes might act in these distinct contexts. The catalytic role of yeast Rcl1 in A2 cleavage was reported to be dependent on an intact RDK motif, which was proposed to mediate RNA substrate binding (9). We found, however, that the $Rcl1_{RDK}$ mutation blocks binding to the SSU processome component Bms1 *in vitro* (Figure 5A), resulting in reduced nuclear import (Figure 5D) and the loss of stable incorporation into the processome (Figure 5B and 5C) *in vivo*. An intact SSU processome is required for A0-A2 cleavage potentially explaining the pre-rRNA processing defect in strains expressing $Rcl1_{RDK}$. Mutation of the equivalent motif in human RCL1 ("RHK") did not affect 18S production (Figure 4). However, human SSU processome assembly appears to differ from yeast, since RCL1 associates prior to A' cleavage, while BMS1 largely binds after cleavage at A' (36,37).

This work has implicated Utp24 as the endonuclease responsible for two coupled pre-rRNA cleavages in 18S rRNA maturation. The enzymes responsible for cleavage at yeast and human site A0 and human A' are probably also among the known SSU processome components, but their identities remain an intriguing mystery.

ACCESSION NUMBERS

GEO database (http://www.ncbi.nlm.nih.gov/geo/): identifier GSE75991.

SUPPLEMENTARY MATERIAL

Supplementary Data are provided as a separate PDF file containing 4 tables, 6 figures and supplementary figure legends.

FUNDING

This work was supported by the Royal Society [UF100666 to C.S., RG110357 to C.S.]; the Wellcome Trust [077248 to D.T., 092076 to N.J.W.]; a BBSRC project grant [BB/F006853/1 to N.J.W]; a BBSRC DTP studentship to [D.C.] and an ERASMUS internship to [F.W.].

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A. and Gleizes, P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA*, **6**, 225-242.
2. Phipps, K.R., Charette, J. and Baserga, S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip Rev RNA*, **2**, 1-21.
3. Watkins, N.J. and Bohnsack, M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA*, **3**, 397-414.
4. Bleichert, F., Granneman, S., Osheim, Y.N., Beyer, A.L. and Baserga, S.J. (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc Natl Acad Sci U S A*, **103**, 9464-9469.
5. Billy, E., Wegierski, T., Nasr, F. and Filipowicz, W. (2000) Rcl1p, the yeast protein similar to the RNA 3'-phosphate cyclase, associates with U3 snoRNP and is required for 18S rRNA biogenesis. *EMBO J*, **19**, 2115-2126.
6. Fatica, A., Oeffinger, M., Dlakic, M. and Tollervey, D. (2003) Nob1p is required for cleavage of the 3' end of 18S rRNA. *Mol Cell Biol*, **23**, 1798-1807.
7. Pertschy, B., Schneider, C., Gnadig, M., Schafer, T., Tollervey, D. and Hurt, E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem*, **284**, 35079-35091.
8. Lebaron, S., Schneider, C., van Nues, R.W., Swiatkowska, A., Walsh, D., Bottcher, B., Granneman, S., Watkins, N.J. and Tollervey, D. (2012) Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat Struct Mol Biol*, **19**, 744-753.
9. Horn, D.M., Mason, S.L. and Karbstein, K. (2011) Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production. *J Biol Chem*, **286**, 34082-34087.
10. Delprato, A., Al Kadri, Y., Perebaskine, N., Monfoulet, C., Henry, Y., Henras, A.K. and Fribourg, S. (2014) Crucial role of the Rcl1p-Bms1p interaction for yeast pre-ribosomal RNA processing. *Nucleic Acids Res*, **42**, 10161-10172.
11. Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J. and Watkins, N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J Cell Biol*, **200**, 577-588.
12. Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L. and Lafontaine, D.L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell*, **51**, 539-551.
13. Sloan, K.E., Bohnsack, M.T., Schneider, C. and Watkins, N.J. (2014) The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA*, **20**, 540-550.
14. Wang, M., Anikin, L. and Pestov, D.G. (2014) Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Res*, **42**, 11180-11191.
15. Tomecki, R., Labno, A., Drazkowska, K., Cysewski, D. and Dziembowski, A. (2015) hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. *RNA biology*, **12**, 1010-1029.
16. Sharma, K. and Tollervey, D. (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Mol Cell Biol*, **19**, 6012-6019.
17. Kos, M. and Tollervey, D. (2010) Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell*, **37**, 809-820.
18. Lazdins, I.B., Delannoy, M. and Sollner-Webb, B. (1997) Analysis of nucleolar transcription and processing domains and pre-rRNA movements by in situ hybridization. *Chromosoma*, **105**, 481-495.

For Peer Review

19. Preti, M., O'Donohue, M.F., Montel-Lehry, N., Bortolin-Cavaille, M.L., Choesmel, V. and Gleizes, P.E. (2013) Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. *Nucleic Acids Res*, **41**, 4709-4723.
20. Gietz, D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res*, **20**, 1425.
21. Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A*, **106**, 9613-9618.
22. Granneman, S., Petfalski, E. and Tollervey, D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J*, **30**, 4006-4019.
23. Wlotzka, W., Kudla, G., Granneman, S. and Tollervey, D. (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J*, **30**, 1790-1803.
24. Schneider, C., Anderson, J.T. and Tollervey, D. (2007) The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol Cell*, **27**, 324-331.
25. Schneider, C., Leung, E., Brown, J. and Tollervey, D. (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res*, **37**, 1127-1140.
26. Schneider, C., Kudla, G., Wlotzka, W., Tuck, A. and Tollervey, D. (2012) Transcriptome-wide analysis of exosome targets. *Mol Cell*, **48**, 422-433.
27. Fayet-Lebaron, E., Atzorn, V., Henry, Y. and Kiss, T. (2009) 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J*, **28**, 1260-1270.
28. Venema, J., Henry, Y. and Tollervey, D. (1995) Two distinct recognition signals define the site of endonucleolytic cleavage at the 5'-end of yeast 18S rRNA. *EMBO J*, **14**, 4883-4892.
29. Rempola, B., Karkusiewicz, I., Piekarska, I. and Rytka, J. (2006) Fcf1p and Fcf2p are novel nucleolar Saccharomyces cerevisiae proteins involved in pre-rRNA processing. *Biochemical and biophysical research communications*, **346**, 546-554.
30. Skruzny, M., Schneider, C., Racz, A., Weng, J., Tollervey, D. and Hurt, E. (2009) An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes. *PLoS biology*, **7**, e8.
31. Allmang, C., Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E. and Tollervey, D. (1996) Recognition of cleavage site A(2) in the yeast pre-rRNA. *RNA*, **2**, 51-62.
32. Elbashir, S.M., Harborth, J., Weber, K. and Tuschl, T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199-213.
33. Wang, M. and Pestov, D.G. (2011) 5'-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. *Nucleic Acids Res*, **39**, 1811-1822.
34. Karbstein, K., Jonas, S. and Doudna, J.A. (2005) An essential GTPase promotes assembly of preribosomal RNA processing complexes. *Mol Cell*, **20**, 633-643.
35. Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., Marcotte, E.M., Tollervey, D., Correll, C.C. and Johnson, A.W. (2015) The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS biology*, **13**, e1002083.
36. Turner, A.J., Knox, A.A., Prieto, J.L., McStay, B. and Watkins, N.J. (2009) A novel small-subunit processome assembly intermediate that contains the U3 snoRNP, nucleolin, RRP5, and DBP4. *Mol Cell Biol*, **29**, 3007-3017.
37. Turner, A.J., Knox, A.A. and Watkins, N.J. (2012) Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA biology*, **9**, 175-186.

For Peer Review

FIGURE LEGENDS

Figure 1: RNA crosslinking sites of yeast Utp24 on (pre-)rRNA and the U3 snoRNA

Illumina sequencing of cDNA libraries generated from crosslinked and subsequently trimmed RNAs recovered with purified Utp24 protein. Normalized data is plotted as reads per million mapped sequences.
A Transcriptome-wide binding profiles. Data of two replicate experiments was mapped to the yeast genome using Novoalign. A total of 979330 mapped reads were recovered for dataset Utp24#1 and 1383819 for dataset Utp24#4, respectively. Pie charts illustrate the proportion of all sequences mapped to functional RNA classes (indicated on the right).
B Sequences from datasets Utp24#1 and Utp24#4 and a control experiment with a non-tagged strain (no tag) were aligned with the rDNA (RDN37-1) encoding the 35S pre-rRNA. The frequency of recovery is plotted as hits (total reads) for each individual nucleotide.
C, E, G RNA binding profiles on the 18S rRNA **(C)**, the 5'-external transcribed spacer (5'-ETS) **(E)** and the U3 snoRNA **(G)**. Hits (upper panels): total reads; dels (lower panels): mutations and microdeletions representing precise binding sites. Microdeletion peaks in the (pre-)rRNA (1-6, blue) and the U3 snoRNA (a-c, red) are labeled. The position of the mature 18S, 5.8S, and 25S rRNAs **(B, C, E)** or the U3 snoRNA **(G)** are indicated by thick lines.
D Predicted secondary structure of the mature 18S rRNA in *S. cerevisiae*. Utp24 crosslinking sites are marked on the sequence and shades indicate peak height with the highest peak shown in dark grey. Microdeletion peaks (see panel **C**) are highlighted by shaded blue circles. Binding sites for snoRNAs U3 around the central pseudoknot (red) and snR30 in the ES6 region (grey) are also indicated.
F The distribution of hits mapping to crosslinked snoRNAs is plotted.
H Schematic representation of base-pairing interactions between the pre-rRNA (black) and the U3 snoRNA (red) during pseudoknot formation. The approximate positions of Utp24 microdeletion peaks are indicated in blue (1-6, pre-rRNA, see panels **C** and **E**) or red (a-c, U3 snoRNA, see panel **G**).

Figure 2: Yeast and human recombinant Utp24 proteins cleave at yeast site A2 *in vitro*

A Schematic of the RNA substrate mimicking a pre-rRNA fragment before A2 cleavage (yeast 35S: nt2301-2844)
B *In vitro* transcribed RNA was incubated without recombinant protein, wild type or mutant Utp24, wild type Rcl1 or wild type or mutant UTP24 in the presence of 5 mM Mn^{2+} and analyzed by primer extension. The position of the primer is shown in panel **A**. Non-treated RNA substrate was used to generate a sequencing ladder and endogenous cleavage at site A2 in yeast total RNA is indicated. Recombinant Utp24-mediated cleavages at site A2 and 5'C/A3' sequences are marked by an arrow and asterisks, respectively.

Figure 3: An intact UTP24 PIN domain is required for efficient site 1 and 2a cleavages in the human pre-ribosomal RNA

A Key steps in human ribosome biogenesis. RNA intermediates of the minor, 2a-dependent pathway accumulating in the absence of XRN2 are shown in red. Radiolabeled probes used for primer extension (h1-RT; orange) or northern blotting (h18SE; purple and hITS1; green) are marked above the 47S precursor. ETS: external transcribed spacer; ITS: internal transcribed spacer.
B HEK293 cells were stably transfected with plasmids expressing the FLAG tag (pcDNA5) or wild type (WT) or mutant forms of FLAG-UTP24 (D72N, D72N/D142N). RNA from control cells (GL2) or those depleted of endogenous UTP24 (UTP24) by RNAi, was analyzed by northern blotting using probes hybridizing to the 5'-end of ITS1 (h18SE, purple rectangle) or downstream of 2a (hITS1, green rectangle). Pre-rRNAs were detected using a PhosphorImager and total rRNA (28S/18S) was visualized by ethidium bromide staining. RNA species are marked on the right.

C RNA extracted from HEK293 cell lines as in panel **B**, but depleted of XRN2 alone (XRN2) or XRN2 and UTP24 (X+24) were analyzed as in **B**.
D RNA levels from panel **C** were normalized to the 47S/45S pre-rRNAs and plotted for each XRN2-UTP24 double knockdown (black) and the single XRN2 knockdown (grey). The identity of each peak is indicated. Red: RNA species accumulating in the absence of XRN2.
E RNA extracted from HEK293 cell expressing the FLAG-UTP24 D72N mutant, either treated with control siRNA (GL2) or depleted of UTP24 (UTP24), XRN2 (XRN2) or both (X+24), was analyzed by primer extension using probe h1-RT (panel **A**). Total RNA from control cells (GL2) expressing the FLAG tag alone (pcDNA5) was used to generate a sequencing ladder. Positions of the natural site 1 and 2 nt downstream are indicated on the left.

Figure 4: **An intact RCL1 RHK domain is not required for human 18S production**

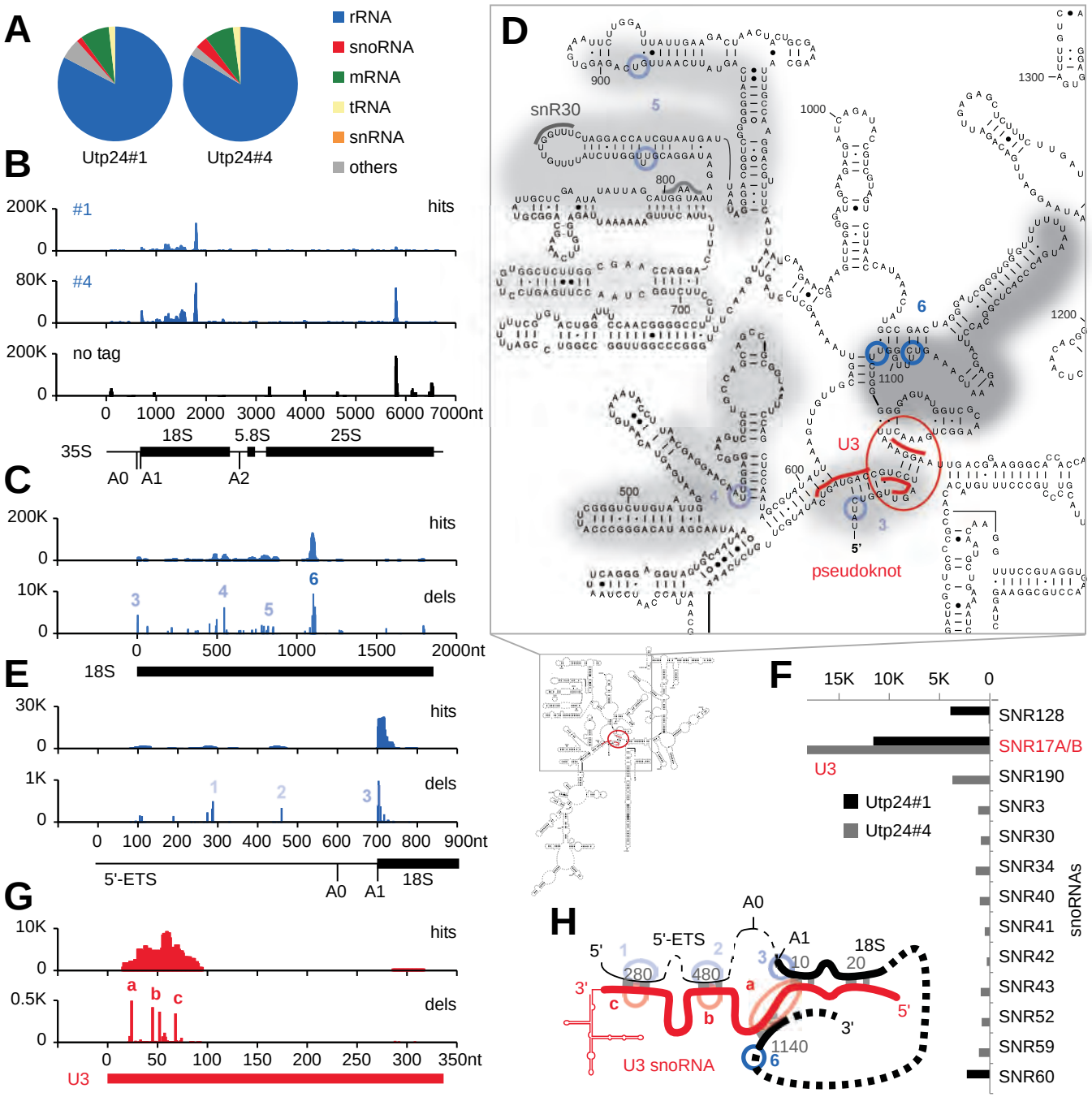
A HEK293 cells were stably transfected with plasmids expressing the FLAG tag alone (pcDNA5) or wild type FLAG-RCL1 (WT) or the FLAG-RCL1-RHK mutant (RHK). The FLAG-RCL1 mRNAs were rendered resistant to the RCL1 siRNA. RNA was extracted from control cells (GL2), or those depleted of endogenous RCL1 (RCL1) by RNAi and analyzed by northern blotting using a probe hybridizing to the 5'-end of ITS1 (h18SE, purple rectangle) or downstream of site 2a (hITS1, green rectangle).
B RNA extracted from cells as listed in panel **A**, but depleted of XRN2 alone (XRN2) or XRN2 and RCL1 (X+R), was analyzed as in **A**.
C The levels of the pre-rRNA intermediates from panel **B** were normalized to the 47S/45S precursors and plotted for each XRN2-RCL1 double knockdown (black) and the single XRN2 knockdown (grey). The identity of each peak is indicated. RNA intermediates accumulating in the absence of XRN2 are shown in red. Asterisk: non-specific signal.

Figure 5: The yeast Rcl1-RDK mutant fails to interact with Bms1 *in vitro* and interferes with SSU processome integration and nucleolar localization *in vivo*.

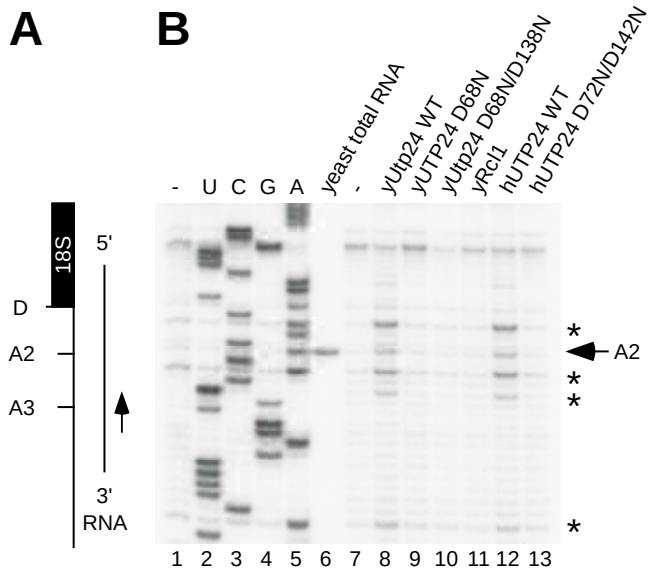
A Recombinant GST-tagged Bms1$^{(aa1-705)}$ and GST were coupled to glutathione sepharose and incubated with wild type and mutant Rcl1 generated by *in vitro* translation in the presence of [^{35}S] methionine. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and analyzed by autoradiography. 10 % of the input and unbound material is also loaded.
B and **C** Rcl1-containing complexes were purified on IgG sepharose and eluted by TEV protease cleavage. Co-purifying proteins were analyzed by SDS-PAGE and immunoblotting using antibodies against the C-terminus of the Rcl1-HTP tag after TEV cleavage (anti-TAP), the box C/D snoRNP component Nop1, or Mtr4 as a control. 2.5 % of the input material is also loaded (**B**). Co-purified RNA was extracted and analyzed by northern blotting using a probe hybridizing to the U3 snoRNA or two regions within the pre-rRNA (yA2-A3, green rectangle or yD-A2, purple rectangle; see **Figure S1A** for location of the probes). Mature 18S and 25S rRNAs were visualized by ethidium bromide staining. 1% of the input material is also loaded.
D Yeast strains expressing plasmid-encoded HTP-tagged wild type Rcl1 (top row) or the Rcl1-RDK mutant (middle row) were harvested and processed for immunofluorescence miscroscopy using the anti-TAP antibody. Lower row: Cells stained with an antibody against endogenous Nop1 as a nucleolar marker. From left to right: Immunofluorescence signal (red), DAPI (blue) and the merged image.

For Peer Review

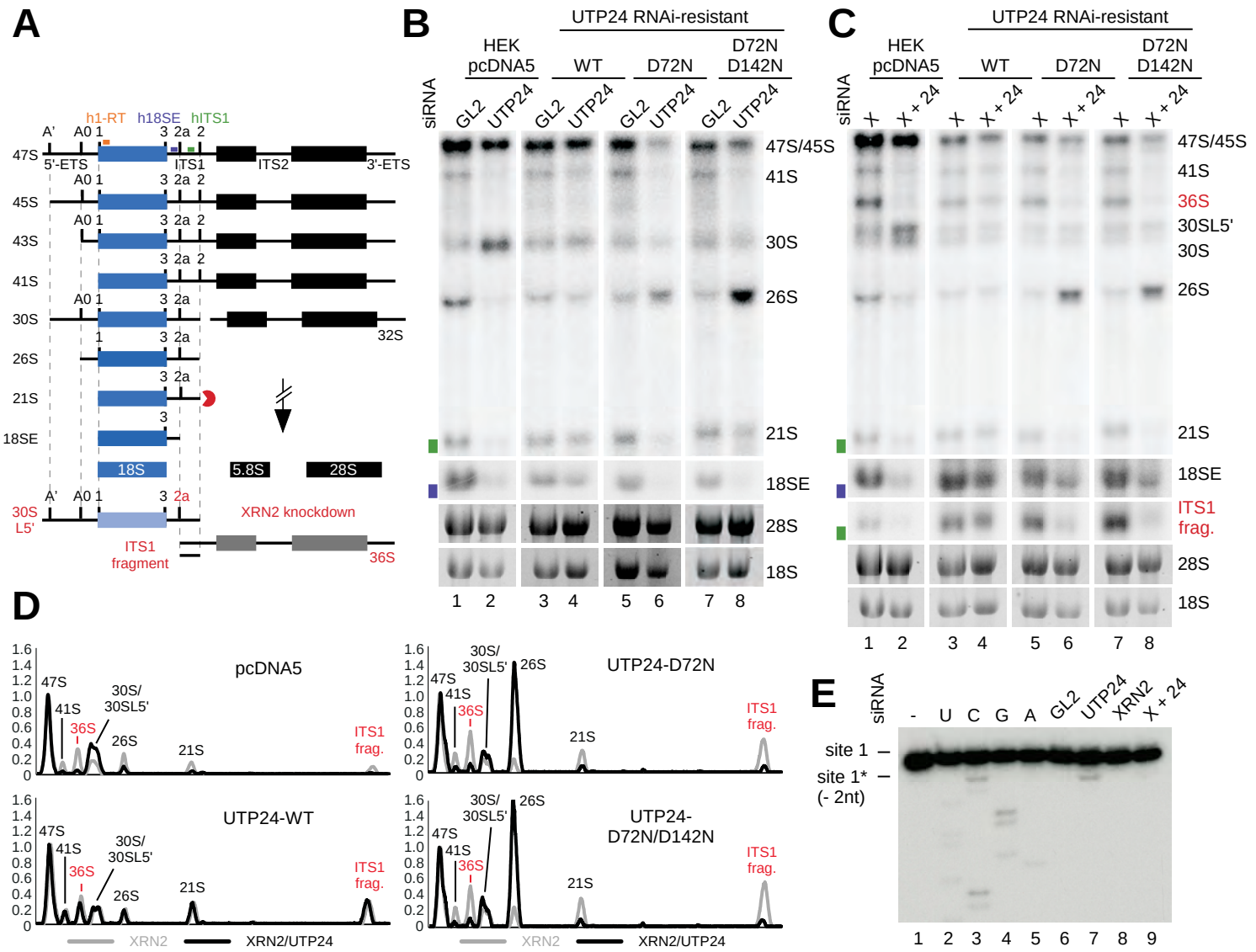
Wells *et al.*, Figure 1



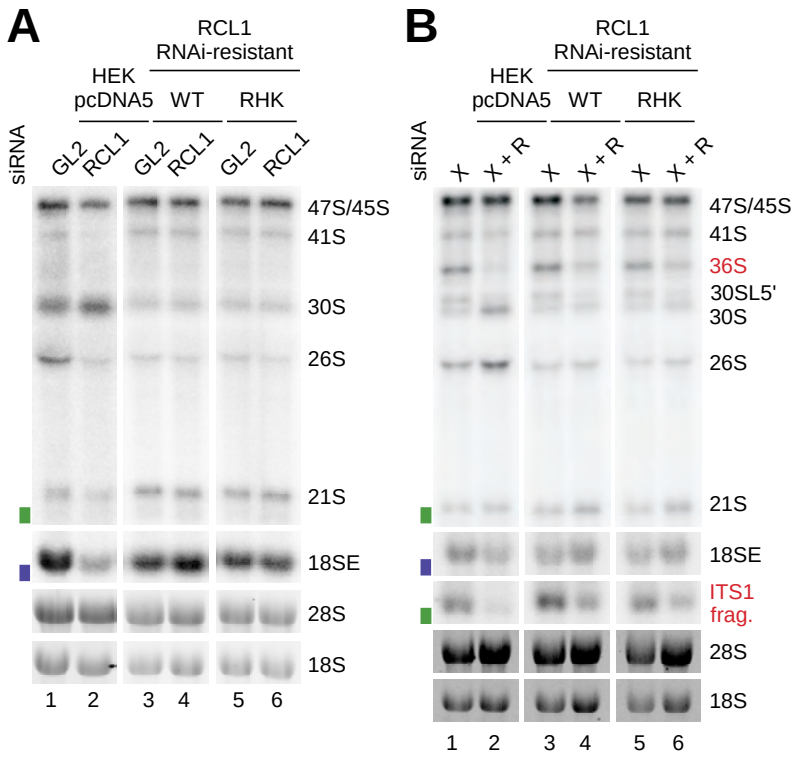
Wells *et al.*, Figure 2

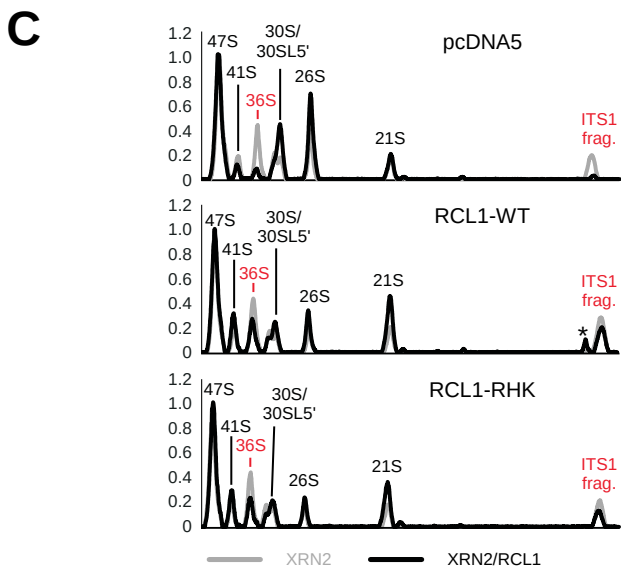


Wells *et al.*, Figure 3

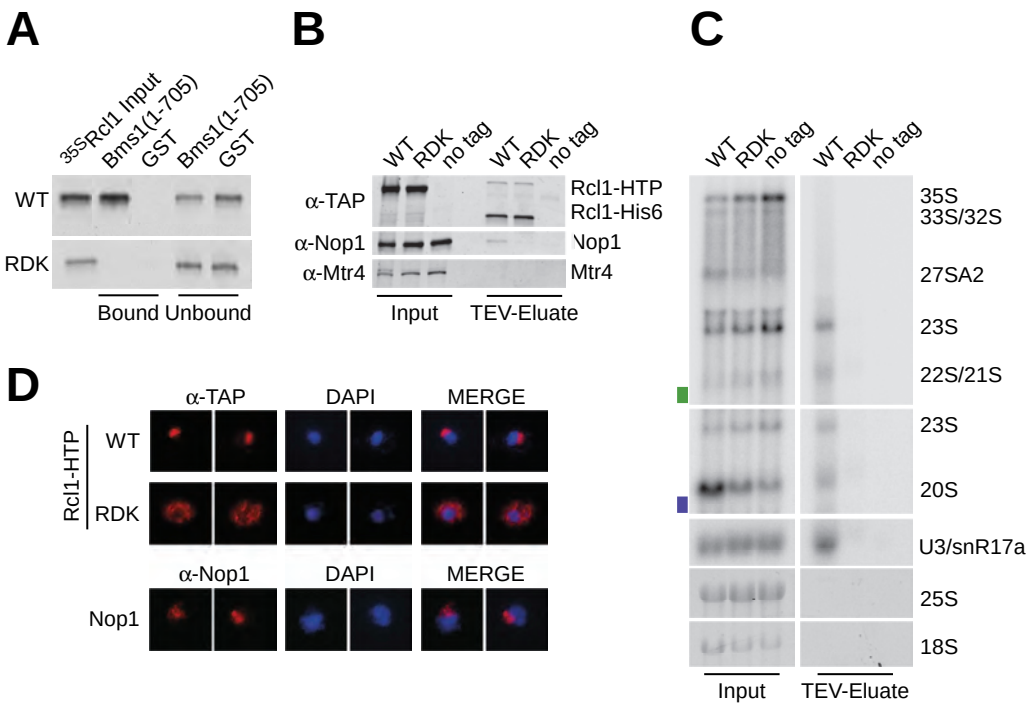


Wells *et al.*, Figure 4





Wells *et al.*, Figure 5



SUPPLEMENTARY TABLES

Table S1: Yeast strains used in these studies

BY4741	MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0	(Euroscarf)
Utp24-HTP	MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; UTP24-HTP-URA3	(this study)
Rcl1-HTP	MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; RCL1-HTP-URA3	(this study)
Gal::rcl1	MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0, KAN::-pGAL1-3HA::rcl1	(this study)

For Peer Review

Table S2: Oligonucleotides (5'-3') used in these studies

Strain construction	
Utp24-HTP	
F	GGCGGTCACGCATACGTCATTGAAAAATTGCCAGATGTCTTTGAG CACCATCACCATCACC
R	CCAAAAACATTGCACTACATACTTAAAGAGTGAATACACAGTAGTT ACGACTCACTATAGGG
Rcl1-HTP	
F	CAAGGGTATTGGTTTCACAAACACAAGCAAAAAGATTGCAGAGCA CCATCACCATCACC
R	ATACAGTGGTCTAATATCTATTGAATAATATTAGAACATGTACGAC TCACTATAGGG
GAL1-3HA::rcl1	
GAL/HA-F4	CATAAGCTACGTATGTTGTAAGGGTATAACATCTGTCAAGGAATTC GAGCTCGTTTAAAC
GAL/HA-R3	ACCCTTGGAAAGTGGTGTATTTGGGGGCGGAAGATGACATGCAC TGAGCAGCGTAATCTG
CRAC	
3'-Linker	5'- rAppTGGAATTCTCGGGTGCCAAGG/ddC/ -3' (miRCat-33^{TM})
5'-Linker (L5Ab)	5'-invddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrU rNrNrNrArUrUrArGrC-OH-3'
RT	CCTTGGCACCCGAGAATT (miRCat-33™)
PCR - F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCT
PCR - R	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGGCC TTGGCACCCGAGAATTCC
Cloning	**(Vector)**
Human	
UTP24	**(pcDNA5 and pGEX-6P1)**
F (Bgl ll)	CACCAGATCTATGGGGAAGCAAAAGAAAACAAGGAAGTATGCGA CCATGAAGCGAATGCTTAG
R (Xho I)	CGCGCTCGAGTTAGAATCGAGGGGCTCCATAATCATCTGG
RCL1	**(pcDNA5)**
F (BamH I)	CACCGGATCCATGGCGACTCAGGCGCACTC
R (Xho I)	CCCCCTCGAGTCACTTGAGGGTCTTGCTAAGGTTGG
Yeast	
Bms1	**(pGEX-6P1)**
F (BamH I)	CACCGGATCCATGGAGCAGTCTAATAAACAGCACCGTAAG
R (Sal I)	GTCGACTTACCTCCTCATCTTACGTGGACGAGATTC
Rcl1	**(pET100 and pGEX-6P1)**
F (BamH I)	CACCGGATCCATGTCATCTTCCGCCCCCAAATAC
R (Sal I)	GTCGACCTATGCAATCTTTTTGCTTGTGTTTGTG
Rcl1-HTP	**(pRS316)**
F (Sal I)	AGCTGTCGACCGAAACATAAAGCTGAAAGACTAG
R (Sal I)	ACGTAGTCGACGCTGGATGGGAAGCGTACC
Utp24	**(pGEX-6P1)**
F (BamH I)	GCGCGGATCCATGGGTAAAGCTAAGAAAACAAGAAAGTTTG
R (Sal I)	CCCCGTCGACTTAAAAGACATCTGGCAATTTTTCAATG
Nob1	**(pGEX-6P1)**
F (Bgl II)	CACCAGATCTATGACCGAAAACCAAACCGCAC
R (Xho I)	CTCGAGCTAACTTCTCCTTTTGGAACTGTTGACGTAC

Mutagenesis	**(Mutation)**
Human	
UTP24	**D72N**
F	CCACCTTACCACATCCTCGTTAATACCAACTTTATCAACTTTTCC
R	GGAAAAGTTGATAAAGTTGGTATTAACGAGGATGTGGTAAGGTGG
UTP24	**D142N**
F	CCATGTACACACAAAGGAACCTATGCAAATGACTGCTTAGTACAG AGAGTAAC
R	GTTACTCTCTGTACTAAGCAGTCATTTGCATAGGTTCCTTTGTGTG TACATGG
RCL1	**RHK -> AAA**
F	CGATAGAATTTTTGGCGGCTTTGGCGAGCTTTTTCCAG
R	CTGGAAAAAGCTCGCCAAAGCCGCCAAAAATTCTATCG
Yeast	
Bms1	**aa705 -> stop**
F	GATAATTCATTCACTAATTTTGATGCGGAGTAAAAAAAGGACTTAA CCATG
R	CATGGTTAAGTCCTTTTTTTACTCCGCATCAAAATTAGTGAATGAA TTATC
Rcl1	**RDK -> AAA**
F	GACGAAAGATTCATAATCCTCTTGGCAGCTATTGCGAAGATCTTTA ATACTGAAGTC
R	GACTTCAGTATTAAAGATCTTCGCAATAGCTGCCAAGAGGATTAT GAATCTTTCGTC
Utp24	**D68N**
F	CAATCAAGCTATAAAGCCACCTTATCAAGTACTGATAAATACCAAT TTTATAAATTTTTCTATCC
R	GGATAGAAAAATTTATAAAATTGGTATTTATCAGTACTTGATAAGG TGGCTTTATAGCTTGATTG
Utp24	**D138N**
F	CGCACAAGGGTACGTACGCGAATGACTGTTTAGTGCATCGAGTC
R	GACTCGATGCACTAAACAGTCATTCGCGTACGTACCCTTGTGCG
Northern blotting	
hITS1 (6121)	AGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC
h18SE	CCTCGCCCTCCGGGCTCCGTTAATGATC
yU3 (snR17a)	ATGGGACTCATCAACCAAGTTGG
yD-A2	CGGTTTTAATTGTCCTA
yA2-A3	TGTTACCTCTGGGCCCCGATTG
Primer extension	
h1-RT	CACTGTACCGGCCGTGCG
yA1-RT	TGAGCCATTCGCAGTTTCAC
yA2-RT	GTTTGTTACCTCTGGGCCCC
yD-RT	CCATCTCTTGTCTTCTTGCCCAG
For transcription	
	(yeast 35S pre-rRNA: nt2301-2844)
y18S-200-T7-F	CGGAATTCTAATACGACTCACTATAGGGCTTGCGTTGATTACGTC CC (T7 promotor)
yITS1-R	CCAGTTACGAAAATTCTTGTTTTTGAC
	(yeast 35S pre-rRNA: nt335-1846)
y35S-nt335-T7F	CGGAATTCTAATACGACTCACTATAGGGAATGCCTTGTTGAATAG CC (T7 promotor)
y35S-nt1846-R	CAATTCCTTTAAGTTTCAGCCTTG

Table S3: siRNAs used in these studies

Gene	Target Sequence (5'-3')	Source
Luciferase GL2	CGUACGCGGAAUACUUCGA	Eurofins MWG (1)
RCL1	GAACAUGACUGUAGCGUCC	Eurofins MWG
UTP24	UCCAAGAUUUGAACGAUUA	Eurofins MWG
XRN2	AAGAGUACAGAUGAUCAUG	Eurofins MWG

For Peer Review

Table S4: Antibodies used in these studies

Target	Source	Cat. No.	Dilution	Species
Western blotting				
α-hKaryopherin	Santa Cruz	sc-11367	1:1000	rabbit
α-hRCL1	Eurogentec (2)	custom	1:250	rabbit
α-hUTP24	Eurogentec (2)	custom	1:1000	rabbit
α-hXRN2	Bethyl Laboratories, Inc.	A301-103A	1: 2000	rabbit
α-TAP*	Thermo Scientific	CAB1001	1:10,000	rabbit
α-yNop1	Santa Cruz	sc-57940	1:2000	mouse
α-yMtr4	Eurogentec	custom	1:10,000	rabbit
α-mouse-HRP	Santa Cruz	sc-2316	1:10,000	donkey
α-rabbit-HRP	Santa Cruz	sc-2313	1:10,000	donkey
Immunofluorescence				
α-TAP	Thermo Scientific	CAB1001	1:1000	rabbit
α-yNop1	Santa Cruz	sc-57940	1:2000	mouse
α-rabbit-IgG (Alexa Fluor 555 conjugate)	Invitrogen	A-31572	1:500	donkey
α-mouse-IgG (Alexa Fluor 555 conjugate)	Invitrogen	A-31570	1:500	donkey

***** The α-TAP antibody recognizes the C-terminus of the HTP construct before and after TEV cleavage.

For Peer Review

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Ribosome biogenesis pathways in yeast and humans

A, B Key steps in the pre-ribosomal RNA processing pathways in *S. cerevisiae* **(A)** and *H. sapiens* **(B)**. Cleavages important for 18S rRNA processing are indicated. RNA intermediates of the minor, 2a-dependent human pathway that are accumulating in the absence of XRN2 (30SL5', 36S and the ITS2 fragment) are highlighted in red. The positions of radiolabeled probes used for primer extension (yA1-RT, yD-RT, yA2-RT and h1-RT) or northern blotting (yD-A2, yA2-A3, h18SE and hITS1) are marked above the primary transcripts. ETS: external transcribed spacer; ITS: internal transcribed spacer.

Figure S2: *In vivo* RNA-protein crosslinking studies (CRAC) to define binding sites for Utp24 and Rcl1

A Outline of the CRAC crosslinking technique.

B Proteins purified from UV-crosslinked yeast strains expressing either C-terminally HTP-tagged Utp24 or Rcl1 were separated by SDS-PAGE and visualized by immunoblotting using the anti-TAP antibody (left panel), which recognizes the C-terminus of the HTP construct after TEV cleavage. Radioactively labeled crosslinked RNA fragments were detected by autoradiography (right panel).

C Mutations and microdeletions on the 5'-end of the 18S rRNA (nucleotides 1-20), representing precise Utp24 binding sites from two replicate experiments, are plotted as deletions per million mapped sequences. The position of the mature 18S rRNA and site A1 are indicated on the left.

Figure S3: Recombinant yeast Utp24 and Nob1 exhibit specific *in vitro* cleavage activities in the presence of 5 mM Mn^{2+} at sites A2 and D, respectively.

A Recombinant, GST-tagged wild type or mutant yeast Utp24, wild type yeast Rcl1 or wild type and mutant human UTP24 proteins were expressed in *E. coli* and purified on a glutathione sepharose column in the presence of 1 mM Mn^{2+}. The N-terminal GST tag was removed by prescission protease (PP) cleavage and the proteins separated by SDS-PAGE and stained with Coomassie blue. Asterisk: free GST-tag.

B *In vitro* transcribed RNA mimicking a 5'-ETS-18S pre-rRNA fragment before A0 and A1 cleavage (yeast 35S pre-rRNA: nt335-1846) was incubated without recombinant protein, wild type or mutant yeast Utp24 protein in the presence of 5 mM Mn^{2+} and analyzed by primer extension. The position of the primer (yA1-RT) is shown. Non-treated RNA substrate was used to generate a sequencing ladder. Recombinant Utp24-mediated cleavages in the 5'-ETS are marked by asterisks.

C Recombinant, GST-tagged wild type yeast Nob1 was expressed in *E. coli*, purified on a glutathione sepharose column in the absence of Mn^{2+} and analyzed by SDS-PAGE and Coomassie staining. Uncleaved, GST-tagged yNob1 protein was used in the cleavage assays shown in panel **D**.

D *In vitro* transcribed RNA mimicking a pre-rRNA fragment before A2 and D site cleavage (yeast 35S pre-rRNA: nt2301-2844) was incubated without recombinant protein, wild type or mutant yeast Utp24 or wild-type Nob1 protein in the presence of 5 mM Mn^{2+} and analyzed by primer extension. The positions of the primers yA2-RT (left) and yD-RT (right) are shown. Non-treated RNA substrate was used to generate a sequencing ladder. Recombinant protein-mediated cleavages at and around sites A2 and D are marked by arrows and asterisks, respectively.

Figure S4: Wild type and mutant recombinant proteins exhibit *in vitro* cleavage activity at site A2 in the presence of 10 mM Mg^{2+}

A Recombinant wild type or mutant yeast Utp24 and wild type yeast Rcl1 proteins as in **Figure S3A** were dialyzed into buffer MKG (50 mM MOPS/KOH pH 7.5, 200 mM KCl, 10 % glycerol, 1 mM DTT, 1 mM THP), and analyzed by SDS-PAGE and Coomassie staining.

B Equal amounts of wild type or mutant yeast Utp24 or wild-type Rcl1 proteins in buffer MKG (see panel **A**), or no protein, were incubated for 1h at 30°C with *in vitro* transcribed RNA mimicking a pre-rRNA fragment before A2 cleavage (yeast 35S pre-rRNA: nt2301-2844). Here, reactions were carried out in the presence of 10 mM Mg^{2+} in 50 mM MOPS/KOH pH 7.5, 100 mM KCl, 2 mM dithiothreitol, 100 µg ml^{-1} bovine serum albumin, 0.8 unit $µl^{-1}$ RNasin, 5% glycerol and 10 ng $µl^{-1}$ *E.coli* tRNA. 10 µl reactions containing ~20 pmol of protein were pre-incubated for 20 min at 30°C before adding 0.125 pmole of the RNA substrate, which was then analyzed by primer extension using primer yA2-RT. Recombinant protein-mediated cleavages at site A2 are marked by an arrow.

Figure S5: Protein and RNA analysis of UTP24 RNAi-rescue cell lines

A HEK293 cells as in **Figure 3** were stably transfected with plasmids expressing the FLAG tag alone (pcDNA5) or wild type (WT) or mutant forms of FLAG-UTP24 (D72N, D72N/D142N). The FLAG-UTP24 cDNA sequence had been modified to render the mRNA resistant to the UTP24 siRNA. Protein extracted from control cells (GL2), or those depleted of endogenous UTP24 (UTP24), XRN2 (XRN2) or both (X+24) by RNAi, was separated by SDS-PAGE and transferred to nitrocellulose membrane. Protein levels were analyzed by immunoblotting using antibodies specific for UTP24 and XRN2, or Karyopherin as loading control. The asterisk denotes a non-specific protein recognized by the antibody.

For Peer Review

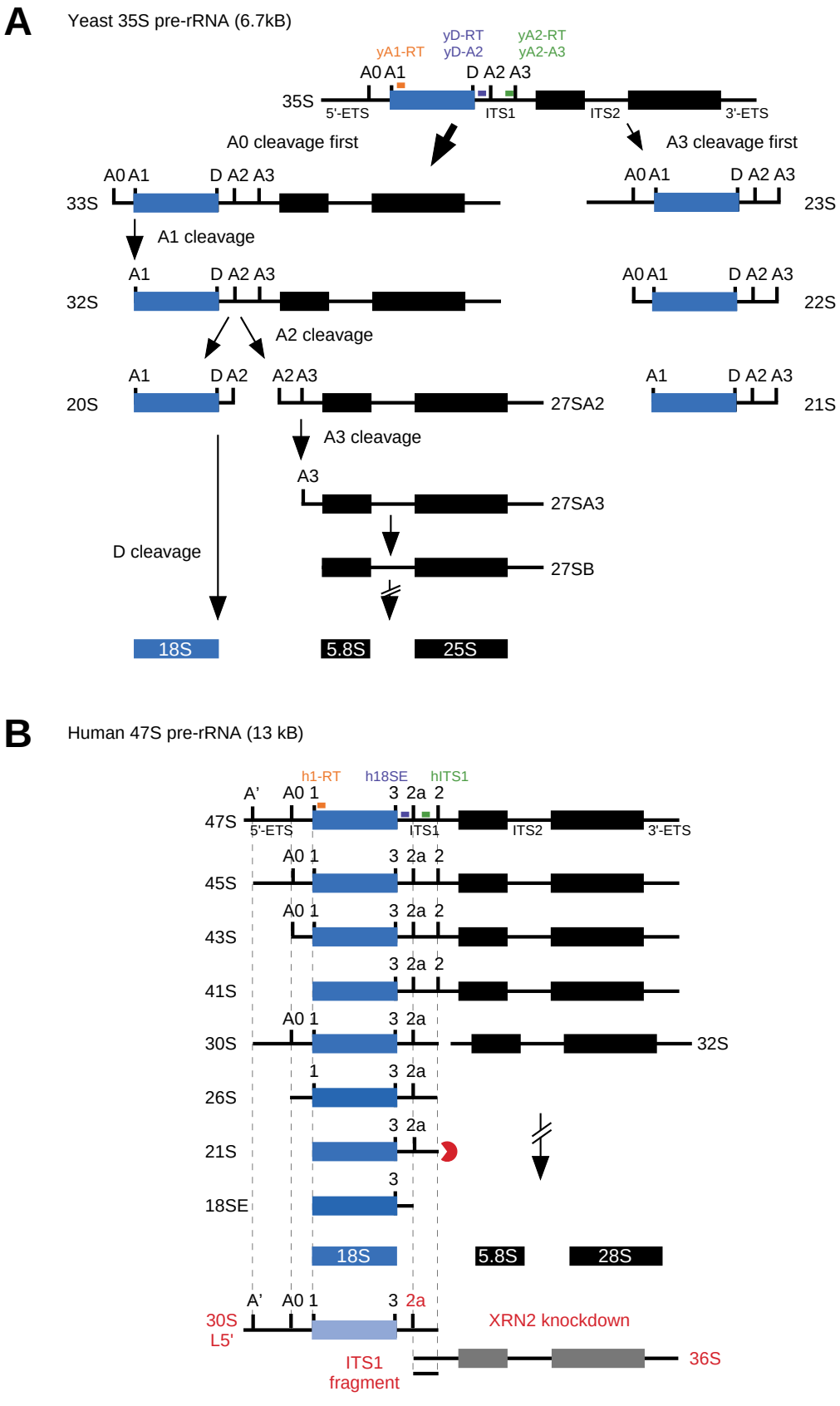
B RNA levels from **Figure 3B** were normalized to the 47S/45S pre-rRNAs and plotted for each GL2 (grey) or UTP24 (black) knockdown. The identity of each peak is indicated.
C RNA extracted from HEK293 cell as in panel **A** was analyzed by primer extension using a probe to detect the 5'-end of 18S (h1-RT). Total RNA from cells expressing the FLAG tag alone (pcDNA5), treated with the control siRNA GL2, was used to generate a sequencing ladder. Positions of the natural site 1 and 2 nt downstream are indicated on the left.

Figure S6: Protein analysis of RCL1 RNAi-rescue cell lines
A HEK293 cells as in **Figure 4** were stably transfected with plasmids expressing the FLAG tag alone (pcDNA5) or wild type FLAG-RCL1 (WT) or the FLAG-RCL1-RHK mutant (RHK). The FLAG-RCL1 mRNAs were rendered resistant to the RCL1 siRNA. Protein extracted from control cells (GL2), or those depleted of endogenous RCL1 (RCL1), XRN2 (XRN2) or both (X+R) by RNAi were analyzed by immunoblotting using antibodies specific for RCL1 and XRN2, or Karyopherin as loading control. The asterisk denotes a non-specific protein recognized by the antibody.
B RNA levels from **Figure 4A** were normalized to the 47S/45S pre-rRNAs and plotted for each GL2 (grey) or RCL1 (black) knockdown. The identity of each peak is indicated.

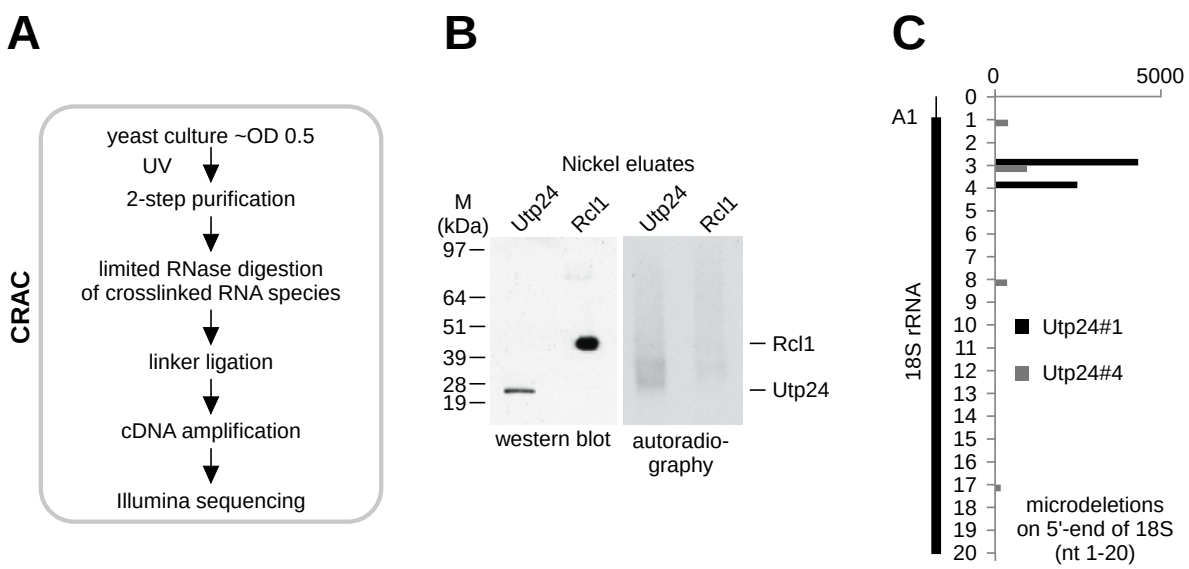
SUPPLEMENTARY REFERENCES

1. Elbashir, S.M., Harborth, J., Weber, K. and Tuschl, T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199-213.
2. Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J. and Watkins, N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J Cell Biol*, **200**, 577-588.

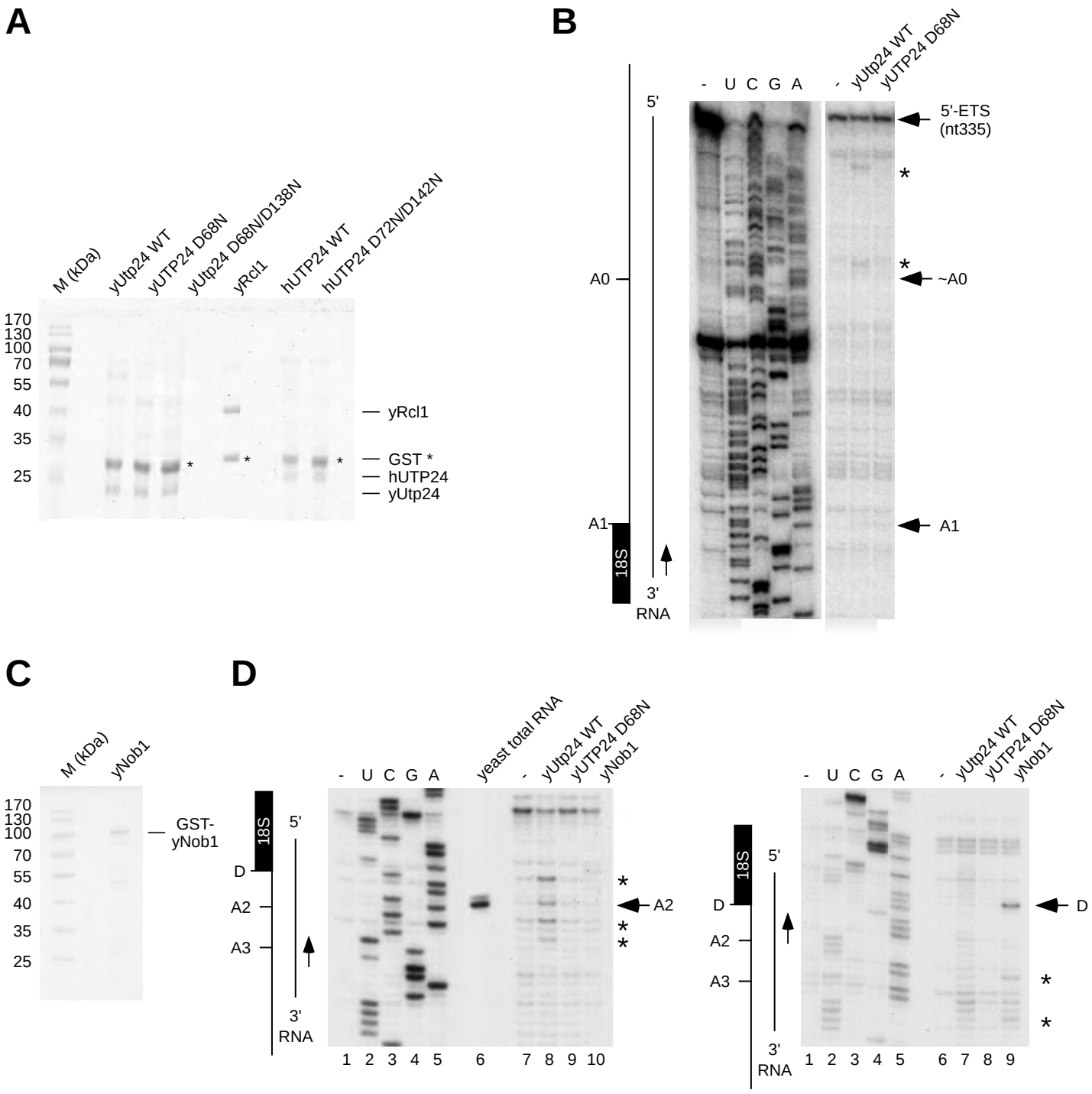
Wells *et al.*, Figure S1



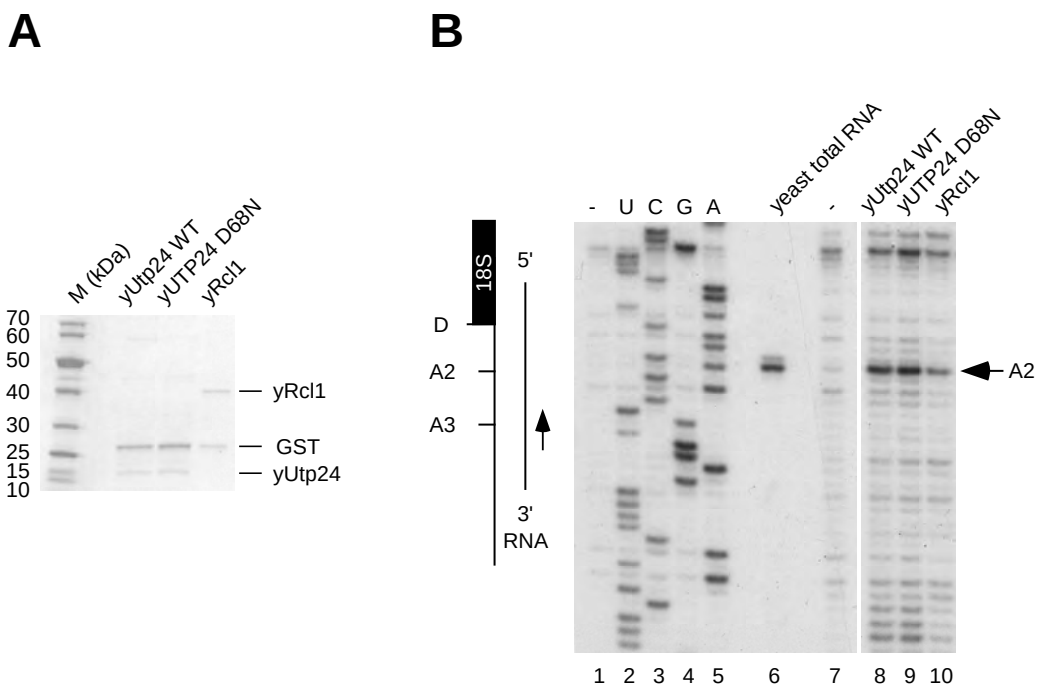
Wells *et al.*, Figure S2



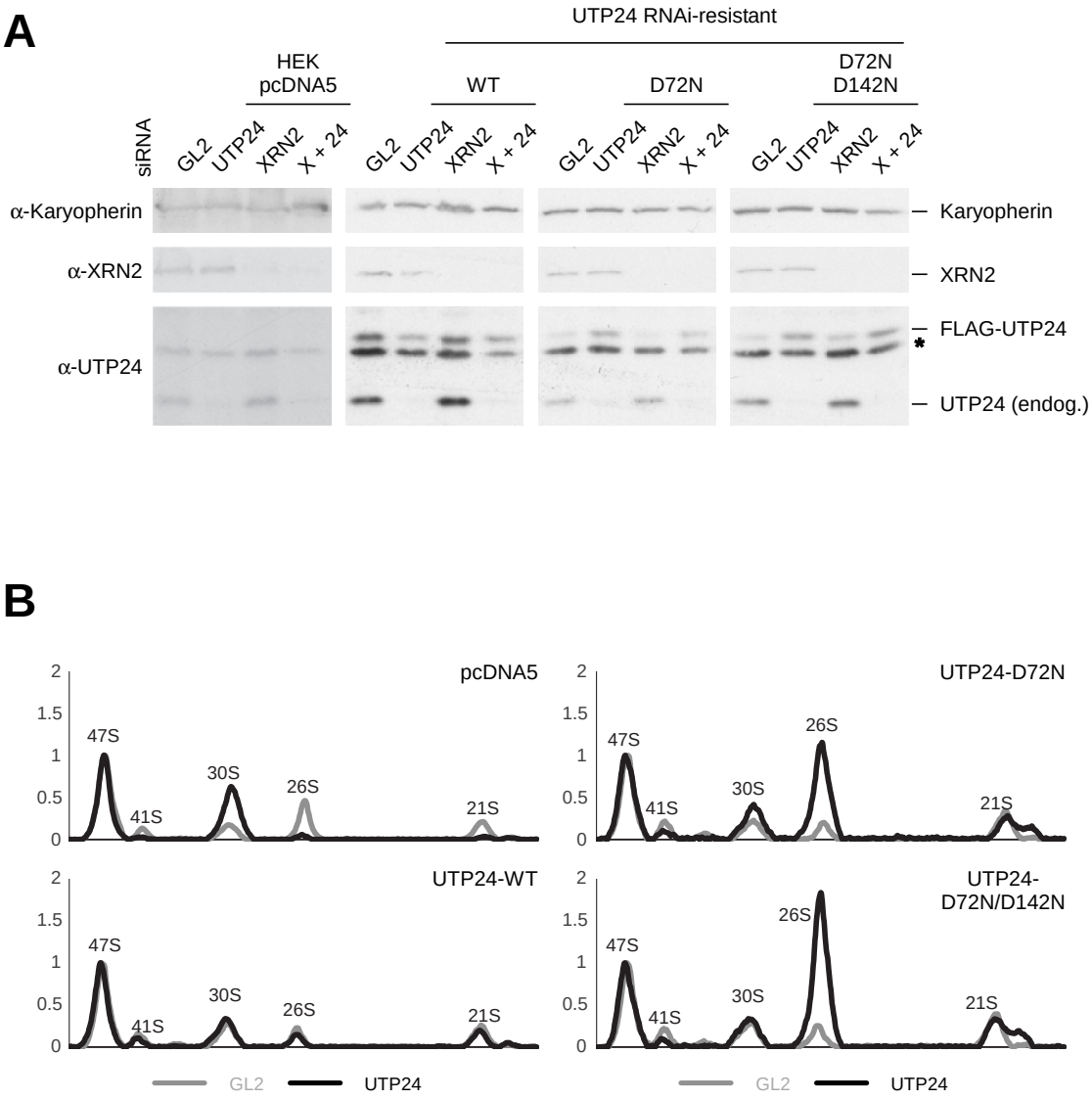
Wells *et al.*, Figure S3



Wells *et al.*, Figure S4



Wells *et al.*, Figure S5



Wells *et al.*, Figure S6